WARM UP:

I have a confidence interval for a mean that is (50.84, 53.76)

(a) What is the sample mean?

(b) What is the margin of error?

(c) Assuming the sample size is 40 and the sample standard deviation is 4.3, what is the level of confidence?

WARM UP:

I have a confidence interval for a mean that is (50.84, 53.76)

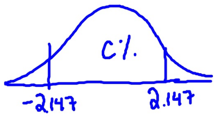
(a) What is the sample mean? $\bar{x} = 52.3$

(b) What is the margin of error? $m = 1.46$

(c) Assuming the sample size is 40 and the sample standard deviation is 4.3, what is the level of confidence?

$1.46 = t^* \left(\frac{4.3}{\sqrt{40}} \right)$

$t^* = 2.147$



$P(-2.147 < t < 2.147) = 96.19\%$

tcdf(-2.147, 2.147, 39) =

Ch. 24: Comparing 2 means

2-sample t-test and interval

2-Sample t-Test

HYPOTHESES:

$H_o: \mu_1 = \mu_2$ OR $\mu_1 - \mu_2 = 0$

$H_a: \mu_1 \gtreqless\neq \mu_2$ OR $\mu_1 - \mu_2 \gtreqless\neq 0$

CONDITIONS:

- just double the 1-sample conditions

(1) 2 independent SRS ① comment on both SRS + indep

(2) $pop_1 \geq 10*n_1$
$pop_2 \geq 10*n_2$ ② comment on both pop

(3) 2 normal populations or
n_1 and $n_2 \geq 30$ ③ check/comment on both

Conditions met => Use t-distribution --> 2-sample t-Test

MECHANICS:

Test statistic:

$$t = \frac{\text{statistic - parameter}}{\text{Std. dev. of statistic (SE)}} \qquad t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

P-value: $P(t \geq \underline{\text{test stat}}) = \text{tcdf}(LB, UB, df)$

on calculator: 2 Samp T-Test

df = on calculator! ***OR*** use the smaller df of the 2 samples

round nearest #

$\boxed{n_1 = 45 \quad df = 44}$ $\quad n_2 = 53 \quad df = 52$

CONCLUSION: (Same 2 sentences)

- reject/fail to reject Ho....
- sufficient/insufficient evidence of Ha

Example: Resting pulse rates for a random sample of 26 smokers had a mean of 80 beats per minute (bpm) and a standard deviation of 5 bpm. Among 32 randomly chosen nonsmokers, the mean and standard deviation were 74 and 6 bpm. Both sets of data were roughly symmetric and had no outliers. Is there evidence of a **difference** in mean pulse rate between smokers and non-smokers?

$n_N = 32$ $n_S = 26$

$\bar{x}_N = 74$ $\bar{x}_S = 80$

$s_N = 6$ $s_S = 5$

Hypotheses:

$H_0: \mu_S = \mu_N$

$H_a: \mu_S \neq \mu_N$

CONDITIONS:

① 2 indep SRS — ① Stated both random samples. Assume smokers indep. of non-smokers.

② $pop_S \geq 10 n_S$, $pop_N \geq 10 n_N$ — ② Assumed more than 260 smokers & 320 nonsmokers.

③ 2 normal pop's or n_1 and $n_2 \geq 30$ — ③ ~~$n_N = 32 \geq 30$~~ Stated both sets of data symm w/ no outliers → assumed normal distrib.

MECHANICS: Cond. met → t distrib → 2 sample t test

$$t = \frac{80 - 74}{\sqrt{\frac{5^2}{26} + \frac{6^2}{32}}} = 4.154$$

$$2P(t > 4.154 \mid df = 56) = 1.129 \times 10^{-4}$$

from calculator when we do 2 samp t test

CONCLUSION:

We reject Ho b/c p-value of $1.129 \times 10^{-4} < \alpha = 0.05$.

We have suff. evid. that the avg. resting pulse rate of smokers is not equal to non-smokers.

2- sample t-Interval:

CONDITIONS: same as 2-sample test

Conditions met --> Use T-distribution --> 2 sample t- Interval

MECHANICS:

statistic ± (critical value)(std. dev. of statistic)

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right) = (a, b)$$

true diff. btw. μ_1 and μ_2

INVT

on calculator: 2 Samp T-Int df = on calculator

CONCLUSION:

We are ____% confident that the difference between the averages of ____ and ____ is between ____ and ____ units.

** Is 0 in the interval???

no diff. btw. μ_1 μ_2

$(-, +)$

Do the confidence interval for our smoking example earlier to estimate the difference. Use 95% confidence.

Conditions met --> t distribution --> 2 sample t Interval
Mechanics:

df = 56
n = 57
INVT

$$(80-74) \pm (2.0032)\sqrt{\frac{5^2}{26}+\frac{6^2}{32}} =$$

$$(3.1063, 8.8937)$$

Conclusion:
We are 95% confident that the difference between the avg resting pulse rate of smokers and non smokers is between 3.1063 bpm and 8.8937 bpm.

Example 2: Here are the saturated fat content (in grams) for several pizzas sold by two national chains. Do the two pizza chains have significantly different mean saturated fat contents? If so, complete an interval to estimate the difference.

L1

Brand D	17	12	10	8	8	10	10	5	16	16
	8	12	15	7	11	11	13	13	11	12
Brand PJ	6	7	11	9	4	4	7	9		
	11	3	4	5	8	5	5			

L2

2 Samp t Test

$\boxed{\text{Data}}$

L1
L2

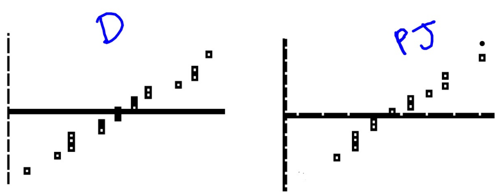
Hypotheses:

$H_0 : \mu_D = \mu_{PJ}$

$H_A : \mu_D \neq \mu_{PJ}$

Conditions:

- 2 independent SRS — assumed representative, & brand D and brand PJ are assumed to be indep.
- pop$_1 \geq 10n_1$, pop$_2 \geq 10n_2$ — there are more than 200 brand D pizzas and 150 brand PJ pizzas sold.
- 2 normal populations or n_1 and $n_2 \geq 30$ — normal probability plots of both samples are approx. linear --> both have norm distributions



Conditions met --> use t-distribution --> 2-samp t-test
Mechanics:

$n_D = 20 \quad \bar{x}_D = 11.25 \quad s_D = 3.193 \quad \alpha = 0.05$

$n_{PJ} = 15 \quad \bar{x}_{PJ} = 6.53 \quad s_{PJ} = 2.588$

df = 33

$$t = \frac{11.25 - 6.533}{\sqrt{\frac{(3.193)^2}{20} + \frac{(2.588)^2}{15}}} = 4.823$$

$2*P(t > 4.823) = 3.151 \times 10^{-5} = 0.0000315$

Conclusion:

We reject Ho b/c p-value of 0.00003 < α = 0.05.

We have sufficient evidence that the average saturated fat content of brand D is not equal to brand PJ.

Interval:

Conditions met --> t-distribution --> 2 sample t interval

df = 33

$$(11.25-6.533) \pm (2.0345)\sqrt{\frac{(3.193)^2}{20} + \frac{(2.588)^2}{15}} = (2.7266, 6.7067)$$

We are 95% confident that the difference between the average saturated fat content of brand PJ and brand D is between 2.7266 and 6.7067 grams.

HW:

p. 579 #2, 3, 11, 14, 27, 36 (a only)